Asthma Stability after Oral Prednisone
A Clinical Model for Comparing Inhaled Steroid Potency

RICHARD C. AHRENS, MARY E. TERESI, SEUNG-HO HAN, DAVID DONNELL, JENNIFER A. VANDEN BURGT, and CHERI R. LUX

The University of Iowa, College of Medicine, Iowa City, Iowa; 3M Pharmaceuticals, St. Paul, Minnesota; and 3M Pharmaceuticals, Loughborough, United Kingdom

Clinical studies comparing the potency of inhaled corticosteroids require steep dose-response slopes (b) and minimal response variability (s), as statistical power is inversely related to the s/b ratio. To evaluate a new study model, we performed a randomized, crossover study of 12 adult asthmatics who required 800 to 2,000 μg of inhaled corticosteroids daily, and calculated s/b for 21 raw clinical outcomes and 36 mathematically derived variables based on these raw outcomes. Each of two 21-d treatment periods was preceded by 4 to 7 d of oral prednisone to maximize asthma control and minimize carry-over of previous inhaled treatment. Treatments were 100 and 800 μ/d of an HFA-134a beclomethasone dipropionate formulation. Assessments included daily home spirometry, histamine challenge, inhaled albuterol use, and asthma symptom scores. Efficacy variables with the greatest power (lowest s/b values) were A.M.FEF_{25-75}, A.M.FEV_1, and A.M.PEF, (s/b = 0.46, 0.48, and 0.59). Carry-over between treatment periods was not significant. Crossover study sample size calculations using these ratios yielded samples of 23, 25, and 37 patients, respectively. Otherwise identical parallel studies would require sample sizes of 657, 1,438, and 2,261 patients. These results support the use of a crossover asthma stability model after a short course of oral prednisone as a clinical study model for comparing topical potency of inhaled corticosteroids.

Keywords: inhaled corticosteroids; potency; outcome variables; sample size; asthma

In the United States, there are currently five corticosteroid compounds available as aerosols for maintenance treatment of asthma (budesonide, beclomethasone dipropionate, flunisolide, fluticasone propionate, and triamcinolone acetonide). These compounds differ with respect to anti-inflammatory potency determined in preclinical studies by the McKenzie skin blanching test, by *in vitro* receptor-binding affinity, and by pharmacokinetic measurement of factors affecting topical potency (1–3).

Accordingly, it is likely that the dose required to produce equal degrees of clinical efficacy also differs between these drugs (4) and even between different formulations of the same drug (5, 6). For example, if 2 μg of hypothetical Formulation A were required to produce the same degree of efficacy as each microgram of Formulation B, Formulation A would be half as potent for producing clinical effect as Formulation B. This relationship between the two formulations, known more formally by the synonyms relative potency, potency ratio, or equal effect dose ratio, is measured using clinical bioassay studies (7, 8). Studies of this type have become an important part of the approval process for generic equivalents to Ventolin and, to a lesser degree, nonchlorofluorocarbon containing beta-adrenergic formulations (9–11). Understanding the potency of different inhaled corticosteroid formulations relative to one another both for topical efficacy and for producing systemic side effects would be useful not only during the drug approval process, but to guide the clinician's choice of specific preparation and dose to prescribe (4).

Unfortunately, accurate comparisons of potency of inhaled corticosteroid formulations relative to one another cannot be obtained from studies published to date (12). Barnes and colleagues (13) concluded that this was due to "shallowness of the inhaled corticosteroid dose-response relationship, high variability of responses obtained, and limitations in the study designs employed" (13).

Dose–response slope, commonly symbolized by statisticians as "b," cannot, in fact, be viewed independently from response variability (e.g., response standard deviation or equivalent, referred to subsequently as "s"). It is the ratio of these two factors (s/b) that determines the statistical power of a clinical bioassay study for estimating the potency of one inhaled corticosteroid formulation relative to another. The smaller the s/b ratio, the more powerful the study (7).

Most investigators attempt to minimize the variability of baseline and response measurements by careful selection of subject inclusion criteria. In spite of this, asthma severity typically varies widely across the patient population studied, as reflected by the level of symptoms, use of inhaled β-agonists, results of spirometry, and measurements of airway responsiveness to histamine or methacholine challenge (14–16). This suggests that use of a crossover study design, which allows each subject to be used as "his or her own control," would substantially increase statistical power.

We postulated that it would be possible to decrease s/b ratio and thereby increase the statistical power of inhaled corticosteroid studies by using a crossover study design and by choosing clinical outcomes with favorable s/b characteristics.

Most investigators have not used crossover designs because of concern about carry-over (5). We sought to control this potential for carry-over corticosteroid effect by beginning each treatment period with a 4 to 7 d course of oral prednisone. In concept, this was intended to equalize carry-over between treatment periods by providing maximal corticosteroid carry-over effect at the beginning of each period. We then monitored *stability* of asthma control during the subsequent inhaled corticosteroid treatment period. This contrasts with most inhaled corticosteroid studies that evaluate *improvement* during inhaled corticosteroid treatment of suboptimally controlled asthma.

Using a two-period, crossover study design, we estimated s/b values for a number of commonly used clinical outcomes. The study was carried out using two dose levels of a new HFA-134a containing beclomethasone preparation (HFA-BDP) developed by 3M Corporation (St. Paul, MN). The purpose was to assess validity of this study design and to choose optimal clinical outcomes to use in subsequent studies estimating

(*Received in original form August 18, 2000; accepted in final form July 5, 2001*)

This was an investigator-initiated project supported in part by a grant from 3M Pharmaceuticals and by Grant RR00059 from the General Clinical Research Center Program, National Center for Research Resources, National Institutes of Health.

Correspondence and requests for reprints should be addressed to Richard C. Ahrens, M.D., Pediatric Allergy/Pulmonary, The University of Iowa, 200 Hawkins Drive, Iowa City, IA 52242-1083. E-mail: richard-ahrens@uiowa.edu

Am J Respir Crit Care Med Vol 164. pp 1138–1145, 2001
Internet address: www.atsjournals.org

the topical potency of HFA-BDP relative to other clinically marketed inhaled corticosteroid formulations.

METHODS

Subjects

Twelve nonsmoking adult subjects with asthma, as defined by the American Thoracic Society (ATS), were enrolled (seven men and five women with a mean age of 27 yr; range, 19 to 39 yr) (17). All could perform spirometry reproducibly as outlined by the ATS standards (18). These subjects had used a stable dose of 800 to 2,000 μg of inhaled BDP, triamcinolone, or fluticasone from a CFC-containing metered-dose inhaler (MDI) for at least 4 wk and used a short-acting β-agonist on an as-needed basis. All had an FEV_1 at screening of greater than 70% predicted, and all had historical evidence of deteriorating asthma when not receiving inhaled corticosteroids. Subjects were excluded if they had had an acute respiratory tract infection within 4 wk of entry; had not used oral corticosteroids at least once within the previous year to treat an acute exacerbation of asthma symptoms; had a history of alcohol or drug abuse; or had used oral β-agonist, cromolyn, nedocromil, salmeterol, or ipratropium within 2 wk of entry, or astemizole within 80 d of entry. The University of Iowa Human Subjects Review Committee approved the protocol, and each subject signed an informed consent before entry into the study.

Study Design

This was a single-center, dose-level blinded, balanced, randomized, two-period crossover study, designed to establish dose-response curves for 100 μg/d and 800 μg/d of HFA BDP. MDIs providing either 50 or 100 μg per actuation (3M) were used to deliver the 100 μg/d (one actuation of 50 μg inhaler twice a day) or 800 μg/d (four actuations of 100 μg inhaler twice a day) of HFA-BDP. Blinding was maintained by use of placebo inhalers, so that subjects always took four actuations twice a day from one inhaler and one actuation twice a day from a second inhaler on each treatment period day.

The study began with a 5- to 14-d run-in period during which subjects continued to take their usual inhaled corticosteroid. This was done to check compliance with data recording. Subjects recorded β-agonist use, sleep disturbance scores, asthma symptom scores and morning (upon arising) and evening spirometry results daily, using a handheld electronic spirometer equipped with daily symptom score recording capabilities (Spiromodem; MultiSpiro, Inc., Irvine, CA). At least three spirometric efforts were obtained at each time point. The highest value at each time point was used in the data analysis. In addition, on even-numbered study days, subjects performed spirometry between 4:00 and 5:00 A.M. (subsequently referred to as 4 A.M. spirometry). Spirometry was also performed if the subject awoke during the night because of asthma symptoms. If this occurred, it was substituted for the 4 A.M. spirometry measurement on even-numbered mornings. On odd numbered days, subjects took two actuations of albuterol immediately after performing morning spirometry and performed a second set of spirometry efforts 15 min later. The subjects were provided albuterol inhalers to use throughout the study. The inhalers were equipped with an electronic dose counter (MDI Chronolog; MEDTRAC Technologies, Inc., Lakewood, CO) that recorded the time and date of each albuterol actuation.

After the run-in period (i.e., just prior to the first treatment period), subjects stopped using their usual corticosteroid inhaler and began a 4- to 7-d course of oral prednisone (40 mg twice a day). Another 4- to 7-d course of prednisone was repeated at the end of the first treatment period (i.e., just prior to the second treatment period). This was done to achieve a maximal corticosteroid effect prior to each inhaled corticosteroid treatment period. Subjects returned to the clinic at the end of prednisone treatment. If the FEV_1 was ⩾ 60% of predicted normal (using an office spirometer: Clinical Pulmonary Function Spirometry System, Med Graphics Corp., St. Paul, MN), a histamine challenge was performed according to a modification of the method of Cockcroft and colleagues (19). A second challenge was initiated 15 min after completion of albuterol administration (180 μg by CFC-MDI). Subjects then began 21 d of study treatment (100 or 800 μg/d HFA-BDP). Data collection during study treatment was the same as noted above for the run-in. Subjects were seen weekly to assess asthma control and to review completion of daily home symptom and spirometry assessments. After 21 d of treatment, a histamine challenge was again performed before and after inhaled albuterol. The

TABLE 1. OUTCOME VARIABLES ANALYZED

	Mathematically Derived Variables							
						Symptom-free (%)		
Raw Variables	% *pred*	Change from Baseline	Change from Baseline (*% pred*)	Change Caused by Albuterol*	A.M./P.M. Variability†	Days	Nights	Both‡
A.M. $FEF_{25-75\%}$, L/s	X	X	X	X	X			
P.M. $FEF_{25-75\%}$, L/s	X	X	X					
4 A.M. $FEF_{25-75\%}$, L/s								
A.M. FEV_1, L	X	X	X	X	X			
P.M. FEV_1, L	X	X	X					
4 A.M. FEV_1, L								
A.M. FVC, L		X			X			
P.M. FVC, L		X						
4 A.M. FVC, L								
A.M. PEF, L/min				X	X			
P.M. PEF, L/min								
4 A.M. PEF, L/min								
A.M. PEF, postalbuterol								
$PC_{20}FEV_1$, prealbuterol‖		X						
$PC_{20}FEV_1$, postalbuterol†§‖		X						
Beta-agonist use, puffs/d		X						X
Sleep disturbance		X					X	X
Wheezing		X				X		
Coughing		X				X		
Chest tightness		X				X		
Shortness of breath		X				X		

* Value measured after two actuations of albuterol MDI, 90μg/actuation minus prealbuterol value.
† A.M. minus P.M. values.
‡ Percent of days and nights free of all asthma symptoms and sleep disturbance.
§ Measured after two actuations of albuterol MDI, 90 μg/actuation.
‖ Provocative concentration of histamine producing a 20% fall in FEV_1

HFA-BDP was stopped, and a second 4- to 7-d course of oral prednisone 40 μg twice a day was started. Treatment with the alternate dose of inhaled HFA-BDP was then administered. Subsequent procedures were the same as those performed during the first 21-d treatment period. Blood hematology and chemistry, ECG, and physical examinations were done during the screening visit and at the final visit.

Outcome Measures

The clinical outcomes measured directly in the study (i.e., raw outcomes) included home spirometry (FEF_{25-75}, [L/s], FEV_1 [L], PEF [L/min], and FVC [L]); histamine PC_{20} FEV_1; asthma symptom and sleep disturbance scores; and the number of actuations of albuterol used per day. Mathematically derived variables consisted of expression of the above raw outcomes in some other manner, e.g., change from baseline (where "baseline" is the value at the end of the prednisone treatment, immediately preceding each treatment period); A.M./P.M. variability (Table 1). For each histamine bronchial provocation, the concentration of histamine that would have produced a 20% decrease in FEV_1 (PC_{20} FEV_1) was estimated by interpolation from a plot of log histamine concentration versus percent decrease in FEV_1 from saline control. The PC_{20} FEV_1 values were log-transformed prior to analysis in order to meet the equal variance assumption of analysis of variance.

Statistical Analysis

The primary analysis utilized the intent-to-treat population. Repeated measures analysis of variance (ANOVA) was used (5% level of significance) to assess the dose-response relation and to test for carry-over effects from the first to the second treatment period. The root mean squared error (s) and dose-response slope (b) values from this ANOVA were used in the sample size computations. Sample size needed to estimate relative potency is a function of the ratio of these two factors (s/b). Slope, b, was defined as:

$$b = (\textit{change in response})/(\textit{change in } \log_{10}[\textit{dose}]).$$

By convention, this slope was given a positive value if the change in response was in the *expected* direction (improvement in asthma with increasing dose). When the change in response went the *wrong* direction (worsening in asthma with increasing dose), the value of b was assumed to be zero. When this occurred, s/b was undefined.

Confidence intervals were estimated for the five variables, with the lowest s/b values using bootstrap analysis (20) with a resampling sample size of 100,000.

Sample size computations assumed administration of two doses of each of the preparations being compared; Finney 2-by-2 bioassay statistical analysis; and 80% probability (power) that the 90% confidence interval for the potency ratio will be less than 2-fold above and greater than 2-fold below the estimate. The method used to calculate sample size took Fieller's theorem into consideration. Sample size computations were made assuming both parallel and complete-block crossover study designs.

RESULTS

Raw outcome variables, without exception, yielded statistical power that was equal to or greater than that of mathematically derived expressions of the same outcomes (i.e., yielded comparable or lower s/b values). Consequently, only data for raw outcome variables are presented here. Measurement of A.M.FEF_{25-75}, A.M.FEV_1, and A.M.PEF produced the lowest values of s/b ratio (0.46, 0.48, 0.59, respectively, for the last 3 d of each treatment period) (Table 2). For each of these three outcomes, the means for the 100 and 800 μg/d doses were similar on Days 1 to 3 of treatment, but they progressively separated as the length of time on treatment increased (Figure 1). Each of the three outcomes showed a highly significant dose-response relationship by Days 19 to 21 ($p = 0.0007$, 0.001, and 0.004).

$PC_{20}FEV_1$ to histamine measured at the end of each treatment period and P.M.FEF_{25-75} measured during the last 3 d of each treatment period were the next most statistically powerful raw outcome variables, with s/b values of 0.81 and 0.98, respectively. The dose-response for these outcomes was also sig-

TABLE 2. RESULTS FOR THE 21 RAW OUTCOMES MEASURED AT THE END OF EACH TREATMENT PERIOD

Outcome Measure	Mean Response 100 μg/d	Mean Response 800 μg/d	Dose-Response Slope* (b)	p-value H_o: Slope = Zero	s†	s/b‡	Sample Size Estimate§
A.M.FEF_{25-75}, L/s	1.80	2.45	0.72	0.0007	0.33	0.46 (0.26, 0.72)	23
A.M.FEV_1, L	2.49	2.85	0.40	0.0010	0.19	0.48 (0.24, 0.72)	25
A.M.PEF, L/min	338.86	379.97	45.52	0.0036	26.66	0.59 (0.27, 1.01)	37
$\log_{10}$, $PC_{20}FEV_1$ (albuterol withheld)	−0.19	0.31	0.55	0.0211	0.44	0.81 (0.16, 2.52)	69
P.M.FEF_{25-75}, L/s	2.00	2.57	0.63	0.0475	0.62	0.98 (0.48, 1.93)	101
A.M.PEF, L/min after two actuations albuterol MDI	418.33	435.25	18.73	0.0850	21.68	1.16	141
P.M.FEV_1, L (albuterol)	2.70	2.88	0.20	0.1620	0.30	1.47	226
4 A.M.FEF_{25-75}, L/s	2.09	2.38	0.33	0.1720	0.50	1.50	238
$\log_{10}$, $PC_{20}FEV_1$ (after two actuations albuterol MDI)	1.09	1.25	0.18	0.1954	0.28	1.59	267
P.M.PEF, L/min	375.63	392.97	19.21	0.2484	34.67	1.81	343
Wheezing score	0.90	0.64	0.29	0.3586	0.67	2.30	556
Actuations albuterol MDI/d, n	3.06	2.61	0.49	0.4919	1.53	3.10	1,010
A.M.FVC, L/min	3.67	3.74	0.08	0.6345	0.37	4.51	2,138
4 A.M.FEV_1, L	2.63	2.68	0.05	0.6991	0.30	5.56	3,247
Sleep disturbance score	0.00	0.17	−0.18	0.0273	0.16	‖	‖
4 A.M.FVC, L	3.63	3.48	−0.17	0.2610	0.31	‖	‖
P.M.FVC, L	3.85	3.73	−0.13	0.4064	0.33	‖	‖
Shortness of breath score	0.89	0.94	−0.06	0.8633	0.77	‖	‖
Coughing score	0.42	0.43	−0.02	0.8987	0.26	‖	‖
4 A.M.PEF, L/min	353.38	352.58	−0.88	0.9433	26.59	‖	‖
Chest tightness score	0.72	0.74	−0.02	0.9538	0.57	‖	‖

* Dose-Response Slope = (Difference in Response)/(Difference in Dose) = (Mean Response for 800 μg/d) − (Mean Response for 100/μg/d)/($\log_{10}$ [800] − $\log_{10}$ [100])

† Root mean square error from ANOVA.

‡ Numbers in parentheses indicate 95% confidence interval for s/b, obtained by bootstrap analysis.

§ Assuming a 2-by-2 crossover bioassay study (high and low dose of each of two inhaled steroid formulations being compared), power of 0.8 and a desired confidence interval around potency ratio estimate with upper limit ≤ twofold above and lower limit ≤ two-fold below estimate).

‖ Indicates s/b undefined since slope ≤ zero (see METHODS).

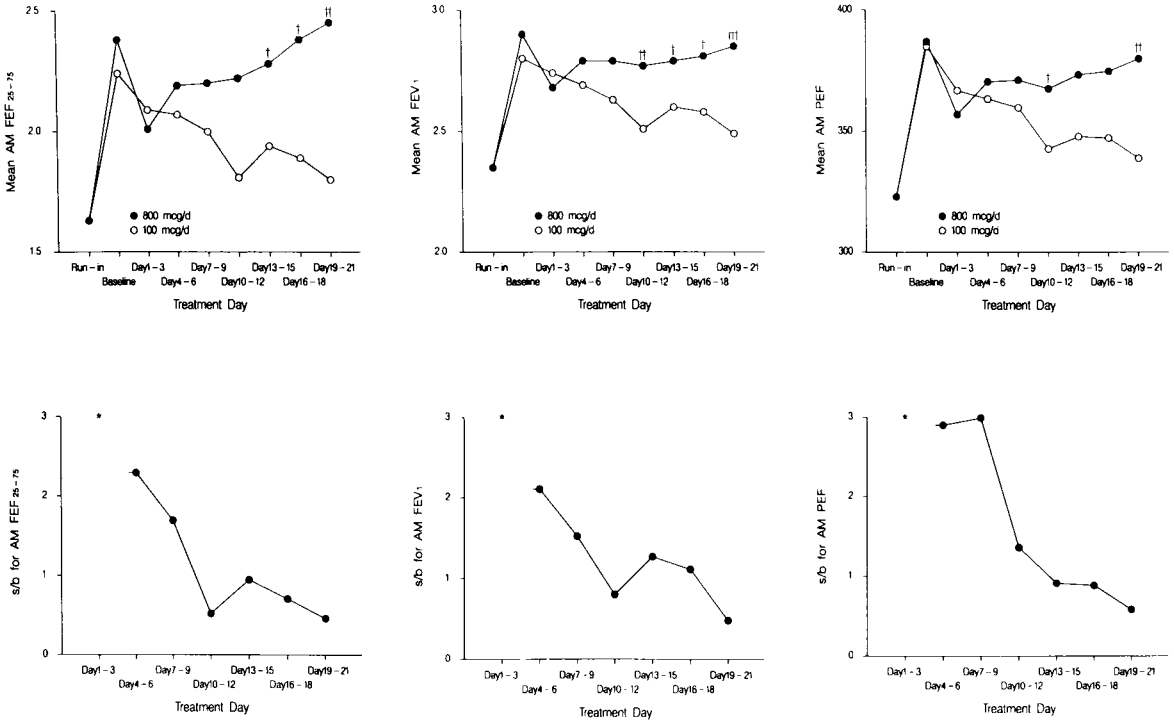

Figure 1. Mean responses by study treatment day for the three outcomes with the lowest s/b values. *Open circles* indicate mean for 100 μg/d of BDP. *Closed circles* indicate mean for 800 μg/d treatment. Difference between treatments by study day: †, ††, ††† indicate $p \leq 0.05$, $p \leq 0.01$, $p \leq 0.001$, respectively. *Asterisk* indicates s/b value undefined because slope = zero (*see* METHODS).

nificant (0.02; 0.048). All other outcome variables had s/b ratios that were > 1.0 or undefined. These outcomes tended to show no consistent pattern of relationship between 100 and 800 μg doses (e.g., mean coughing symptom score and 4 A.M. PEF) (Figure 2).

Treatment sequence (i.e., whether the 100- or the 800-μg

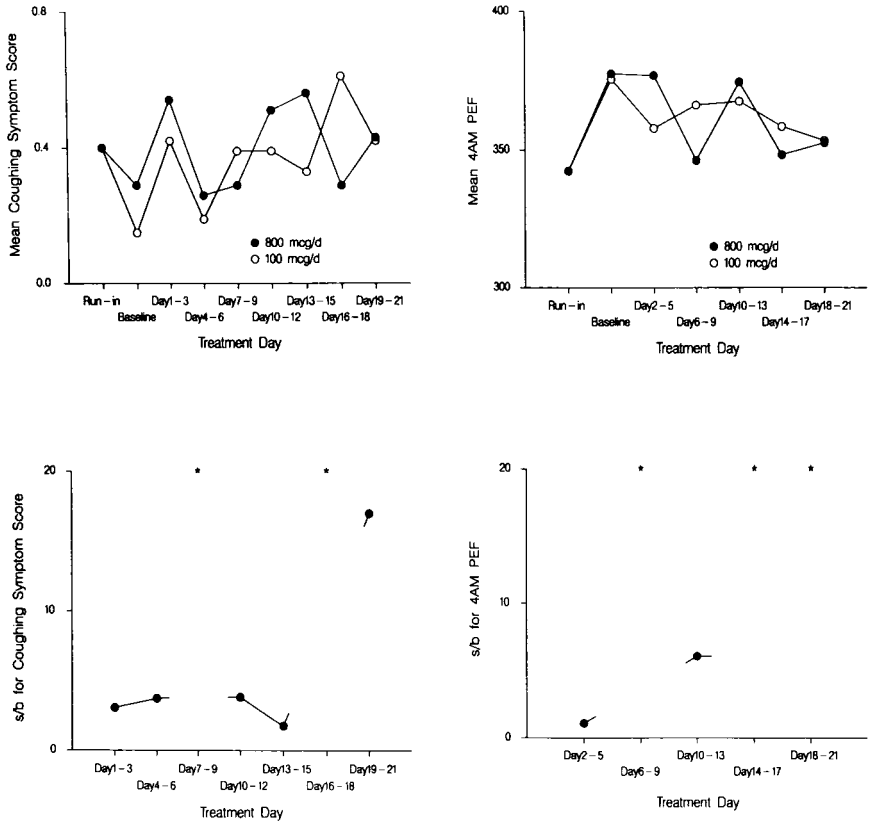

Figure 2. Mean responses by study treatment day for two of the outcomes with the highest s/b values. *Open circles* indicate mean for 100 μg/d of BDP. *Closed circles* indicate mean for 800 μg/d treatment. *Asterisks* indicate points at which s/b values are undefined because slope = zero (*see* METHODS).

dose was given in the first treatment period) was not significant for any of these five best outcomes. This indicates that there was no evidence of carry-over of effect from the first to the second treatment period for any of these outcomes.

Sample size calculations show that for the least powerful outcome variables for which s/b could be defined (see METHODS), thousands of subjects would be required (Figure 3, Panel A). The most powerful outcomes (A.M.FEF_{25-75}, A.M.FEV_1, and A.M. PEF_{25-75}) can be expected to allow clinically useful estimates of relative potency with less than 50 subjects (Figure 3, Panel C).

Sample size calculations made using data from this study, assuming (*1*) parallel and (*2*) crossover study design, show that for the same outcomes, far fewer subjects are required for crossover designs. Even the most successful published attempt to determine a potency ratio for inhaled corticosteroids using a parallel design (Busse and colleagues, 6) would require more than 700 subjects in order to provide a similarly precise estimate of relative potency (Figure 4).

Adverse Events

During each study drug treatment period, 10 (83.3%) subjects reported at least one adverse event. Reported events most commonly involved the respiratory system. The most common respiratory event was increased asthma symptoms, reported by three (25%) subjects while receiving 100 μg HFA-BDP and three subjects while receiving 800 μg HFA-BDP. Pharyngitis occurred in four (33%) subjects receiving 100 μg HFA-BDP and in two subjects (16.7%) receiving 800 μg HFA-BDP. There were no serious or clinically important adverse events or changes in laboratory values during the study.

DISCUSSION

Our results indicate that use of stability of asthma control after a short course of oral prednisone study model, raw A.M.FEV_1, A.M.FEF_{25-75}, or A.M.PEF as the primary outcome, and a crossover study design will allow accurate, precise, and clinically relevant comparisons of inhaled corticosteroid potency to be made, whereas to date, this has not been possible. There were broad differences in statistical power associated with the outcome variables we evaluated. Raw outcome variables (e.g., A.M.FEV_1, A.M.PEF) were consistently associated with equal or greater statistical power than mathematically derived outcomes (e.g., change from baseline A.M.FEV_1, A.M. to P.M. peak flow variability). Some of these raw outcomes were much more powerful than others, with spirometry measured daily upon arising providing the greatest statistical power (A.M.FEF_{25-75}, A.M.FEV_1, A.M.PEF). In contrast, statistical power associated with other raw outcomes appeared to be so low that it could not be defined (Table 2). Perhaps the most important implication of the data presented here is that use of crossover rather than parallel design is essential if a study is to accurately and precisely compare the potency of different inhaled corticosteroid formulations. Concern about carry-over of inhaled corticosteroid effect from one treatment period to the subsequent period has commonly been used as an argument against crossover designs for comparing inhaled corticosteroid potency (5). The study design we used, stability of asthma control after a short course of oral prednisone, appears to accomplish the intended purpose of controlling carry-over in that no statistically significant carry-over was present.

It is not unexpected that some outcome variables are associated with greater statistical power than are others, as indicated by lower s/b values. We have previously presented similar findings for outcome variables used in the comparison of potency of different inhaled albuterol formulations (9, 10). It

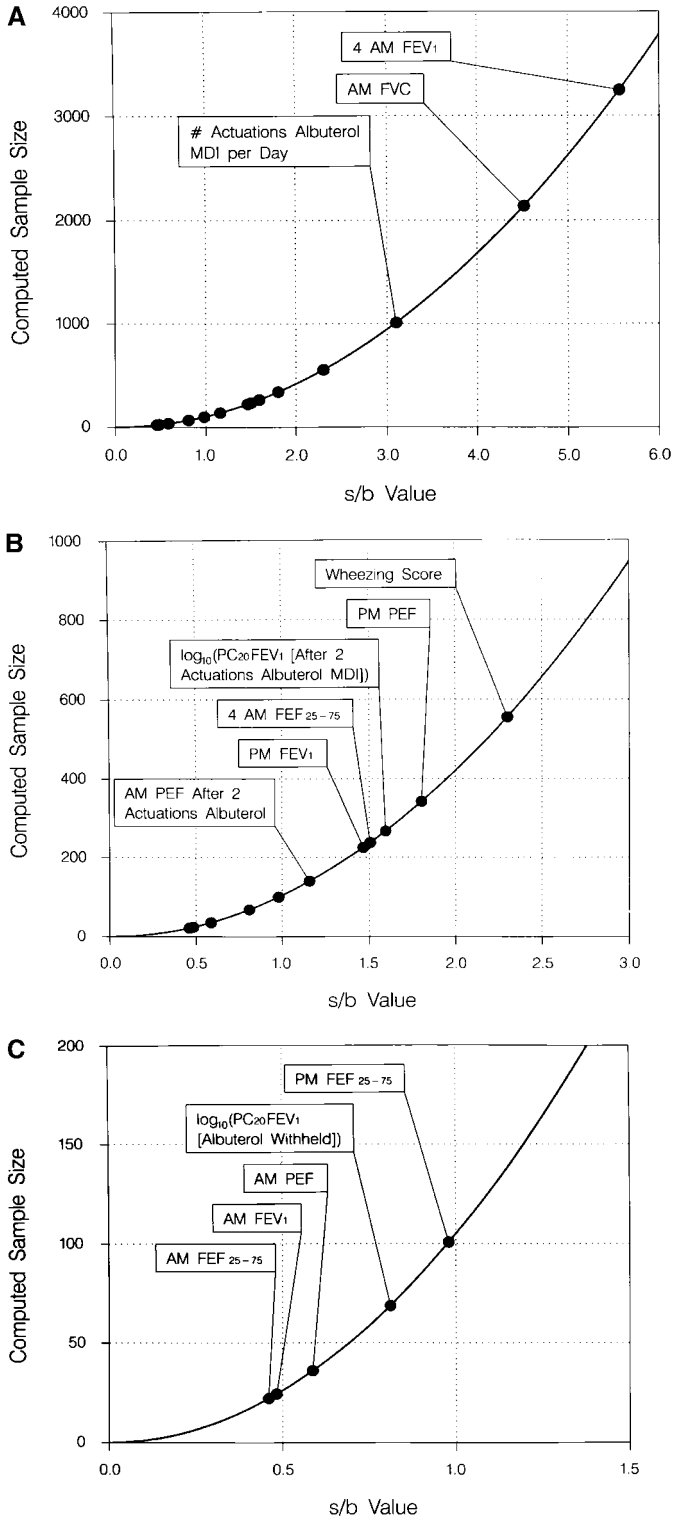

Figure 3. Relationship between s/b value and sample size for 2-by-2 (4-period) crossover bioassay design, power of 0.8, and a desired confidence interval around potency ratio with an upper limit ⩽ 2-fold above and lower limit ⩽ 2-fold below the potency ratio estimate. (*A*) Shows labels for the least power outcome for which s/b could be defined. (*B* and *C*) Sequential "zoom in" on the lower left hand corner of (*A*), showing labels for increasingly more powerful outcomes (i.e., outcomes requiring progressively smaller sample sizes).

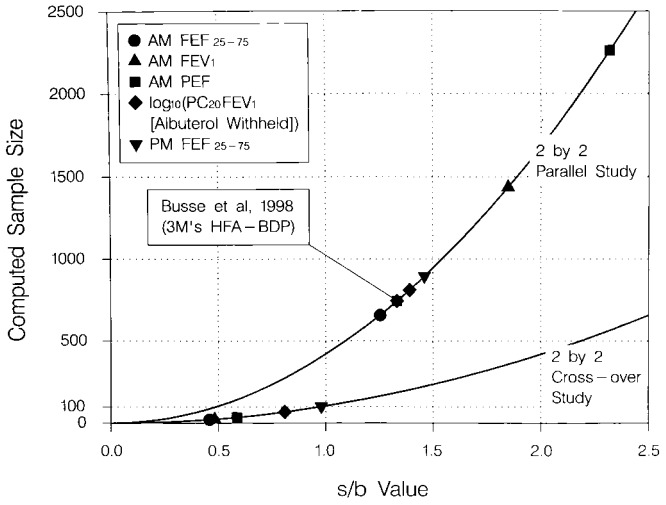

Figure 4. Relationship between sample sizes required for parallel versus four-period crossover study designs. Sample sizes for both designs were computed using data from the current study and are consistently much larger for parallel study designs. For comparison, sample sizes based on a published parallel study (Busse and colleagues, 6) estimating the potency of an HFA containing BDP formulation (3M Pharmaceuticals) relative to a CFC containing BDP formulation (Schering, Inc.) using A.M.FEV_1 is also plotted.

would, of course, be desirable to identify even more powerful outcomes. However, statistical power is only one of three factors to be considered when choosing a clinical model and outcome variables for an inhaled corticosteroid study. Ease and practicality of performance of the measurement and clinical relevance are also important. Outcome measures that are labor-intensive and require specialized laboratories and highly trained personnel are less desirable than are those that do not. Study designs employing models that mimic typical patterns of clinical use and outcomes that are commonly measured in the office or clinic are considered by some to be more clinically relevant than those that do not.

The model we have presented here, stability of asthma after a short course of oral corticosteroid, fairs well for each of these three factors. Not only does it provide sufficient statistical power, it emulates a measurement that is in common clinical practice, daily recording of PEF. Our use of an electronic home spirometer rather than a plastic peak flow meter builds upon this. It allows measurement of FEV_1 and FEF_{25-75} in addition to PEF, and avoids the possibility of fabrication of paper-recorded data by time and date stamping of electronically recorded data. This study model is clinically relevant in that it mimics a common clinical practice: use of oral corticosteroids to reverse existing asthma symptoms and airway obstruction prior to initiating inhaled corticosteroid treatment intended to maintain asthma control. This approach is recommended in the current NIH asthma treatment guidelines (4). In contrast, most studies evaluating inhaled corticosteroid efficacy measure the reversal of airway obstruction and asthma symptoms by the inhaled corticosteroid from a baseline of poorly controlled asthma (5, 13).

As part of the search for more precise, statistically powerful ways to compare inhaled corticosteroids, a number of other outcomes have been evaluated by other investigators. Possibilities include measurement of inhaled corticosteroid effects on sputum eosinophilia, exhaled nitric oxide (NO), allergen challenge, methacholine challenge, and adenosine challenge. Although each of these measures has been shown to respond to corticosteroid treatment (21–26) only limited data are available upon which to judge their statistical power for estimating inhaled steroid relative potency. Available publications have not presented the values for s/b needed for sample size computation. Sputum eosinophilia appears to demonstrate a dose-response relationship throughout the clinically relevant dose range (27) but variability appears to be large (23). Exhaled NO responses appear to plateau at the top of the dose-response curves at relatively low inhaled corticosteroid doses (27). This suggests that the s/b values for both sputum eosinophilia and exhaled NO may be unfavorable (i.e., large). Histamine and methacholine challenge may be relatively powerful for establishing dose-response relationships and for estimating relative potency of inhaled steroids based on the current study, and the work of Kraan and colleagues (28). Others have failed to find significant dose-response relationships for inhaled corticosteroid effects on histamine and methacholine challenge, but this may be due to a treatment duration of only 1 wk (25), and inclusion of patients with only mild asthma (29). Both adenosine and especially exercise challenge appear to be statistically powerful (24, 30) and capable of establishing highly significant dose-response relationships in studies with only small numbers of subjects. None of these outcomes are as convenient and practical as home spirometry measurement. All require specialized laboratories and trained personnel. This is particularly true for sputum eosinophilia where even in experienced hands, as much as 20% of the patients can not produce enough sputum to be analyzed (21, 22). All of the above measures have clinical relevance, although the relative merit of each is a common source of debate.

Our finding that raw outcomes consistently provided greater statistical power for crossover studies than did derived outcomes such as change from baseline, runs contrary to the intuition of many investigators (many inhaled corticosteroid comparison trials have used change from baseline). Subtracting the baseline value actually increases variance and thereby decreases statistical power, *unless* the response to treatment is strongly correlated with the baseline value (31). In the absence of correlation, when one variable (e.g., pretreatment baseline) is subtracted from another (e.g., the posttreatment response), the variance of the resultant difference will be equal to the sum of the variances of these two variables. This will decrease rather than increase statistical power. However, under the right conditions, use of change from baseline can and often does increase statistical power. This occurs when a component of variability in the measured response to treatment is due to variability in the baseline value measured in that patient (i.e., when the posttreatment response is correlated with the pretreatment baseline value). Subtracting the baseline is one method of removing this variability from the measured response, thereby tending to increase power. If the correlation is strong enough, it can more than compensate for the additivity of baseline and posttreatment variance noted above (32). The net effect will be a decrease in variance and an increase in power. For the study model explored here, it seems clear that the correlation between baseline and posttreatment response values is insufficient to overcome the loss in power caused by additivity of baseline and posttreatment variance.

The greater statistical power associated with crossover, as opposed to parallel study designs is well recognized. What does not appear to have been recognized previously is just how important this is to inhaled corticosteroid comparison studies. Published studies comparing inhaled corticosteroids nearly always use a parallel design and rarely provide reliable information about the relative potency of the formulations being compared. Only a single published study has succeeded in estimating a relative potency for different inhaled corticosteroid prepara-

tions. Busse and colleagues (6) estimated that each actuation or microgram of HFA-BDP (QVAR; 3M Pharmaceuticals) was 2.6 times as potent as each microgram or actuation of CFC-BDP (Vanceril; Schering, Inc., Kenilworth, NJ) for improving FEV_1. Even though that study was among the most rigorously designed and conducted study ever carried out, the confidence intervals on these estimates were too wide to be of clinical use (1.1 to 11.6). To achieve a confidence interval sufficiently narrow to be useful clinically (e.g., no more than twofold above and below the potency ratio estimate), sample size for that study would have been at least 700 (Figure 4). Similarly, Figure 4 demonstrates that if the outcomes identified in the current study as being most statistically powerful (A.M. spirometry) were used in an otherwise identical parallel study, the sample size needed would increase by an order of magnitude. This suggests that accurate comparisons of inhaled corticosteroid potency will be more practical and, in fact, may only be possible when done using crossover study designs.

The much greater statistical power associated with crossover inhaled corticosteroid studies in itself suggests that the inhaled corticosteroid dose-response curve is not as flat as many investigators believe (at least not in the population of patients included in the current study). We found a highly significant dose-response "signal" for morning (A.M.) spirometry variables in our 12-patient crossover study. In contrast, in parallel studies, the dose-response "signal" appears to get lost in the between-patient variability—thus making the dose-response relationship seem to be flatter than it really is. This may be, at least in part, because some patients do reach the top of their dose-response relationship with the lowest dose of inhaled corticosteroid, whereas others are on the steeply rising portion of the curve through the range of doses typically used clinically. The current study selectively included patients on the rising portion of the curve based on a clinical history of deterioration of asthma control when inhaled corticosteroids dose was reduced. Future studies will need to have entrance criteria that similarly select for this clinically important group of patients.

The importance of using crossover designs for establishing dose–response relationship and comparing potency of inhaled corticosteroid is supported by previous publications. Pedersen and Hansen (30), using response to exercise challenge, found a significant relationship for inhaled budesonide ($p < 0.0001$) in a crossover study of only 19 children with asthma. Johansson and Dahl (32), using PEF, found a significant relationship for inhaled budesonide ($p < 0.05$) in a crossover study of eight adults with asthma. Taylor and colleagues (24), using adenosine, found a significant relationship for inhaled ciclesonide ($p = 0.04$) in a crossover study of 29 asthmatic adults. In contrast, parallel studies enrolling hundreds of subjects often have failed to find a significant dose-response.

If crossover designs are to be used, the potential for carry-over of effects between study treatment periods must be controlled. Our study suggests that this is possible since we did not detect any carry-over. It is possible we did not have sufficient statistical power to detect a small carry-over effect. However, even if some carry-over does exist, this may not exclude the use of crossover designs, with proper study design, carry-over can be accounted for statistically so as to not bias comparisons of drug potency (33).

Sample size calculations presented here (Figure 3) assume a four-period crossover design (4 wk per period) during which two doses of each formulation are given. In reality, this study may be too long to be practical. However, use of a balanced incomplete block design, where each subject receives two or three of the possible four treatments, would make the study more practical, retain much of the advantage of the crossover design (increased power), and only modestly increase the number of subjects needed.

Choice of doses for each of the formulations studied will also be important for the design of future studies. Both high and low doses should ideally be on the steep portion of the dose-response relationship. If the lower doses are so low as to have little effect or if the higher doses approach the maximal response, dose-response slope will be lowered. As a result, precision and statistical power of the study would be diminished, and the confidence interval around the potency ratio would be widened.

In summary, we have evaluated a new model for comparison of efficacy of inhaled corticosteroid preparations: stability of asthma control after a short course of oral prednisone. The essential features of this model are crossover design, initiation of each study period with oral prednisone, and measurement of raw daily A.M. spirometry as the primary outcome during subsequent inhaled corticosteroid study treatments. This model appears to provide the statistical power needed to estimate relative potency of different inhaled corticosteroid formulations with 90% confidence intervals within twofold above and below this estimate with enrollment of less than 100 subjects. This model is practical and clinically relevant. The potential utility of this model needs to be confirmed with subsequent studies that compare the potency of two different inhaled corticosteroid formulations.

Acknowledgment: The writers would like to thank Donna Reihman for clerical work in preparation of the manuscript.

References

1. Johnson M. Pharmacodynamics and pharmacokinetics of inhaled glucocorticoids. *J Allergy Clin Immunol* 1996;97:169–176.
2. Brattsand R, Thalen A, Roempke K, Kallstrom L, Gruvstad E. Development of new glucocorticoids with a very highratio between topical and systemic activities. *Eur J Respir Dis* 1982;63(Suppl:62–73).
3. Dahlberg E, Thalen A, Brattsand R, Gustafsson JA, Johansson U, Roempke K, Saartok T. Correlation between chemical structure receptor binding and biological activity of some novel, highly active 16 alpha, 17 alpha-acetal-substituted glucocorticoids. *Mol Pharmacol* 1984;25:70–78.
4. National Institutes of Health. Expert Panel Report 2. Guidelines for the Diagnosis and Management of Asthma. Clinical Practice Guidelines. NIH Publication No. 97-4051. April 1997.
5. Kelly HW. Comparative potency and clinical efficacy of inhaled corticosteroids. *Respir Care Clin North Am* 1999;5:537–553.
6. Busse WW, Brazinsky S, Jacobson K, Stricker W, Schmitt K, Vanden Burgt J, Donnell S, Hannon S, Colice GL. Efficacy response of inhaled beclomethasone dipropionate in asthma is proportional to dose and is improved by formulation with a new propellant. *J Allergy Clin Immunol* 1999;104:1215–1222.
7. Finney DJ. Statistical Methods on Biological Assay. 3rd ed. New York: MacMillan Co.; 1978. p. 1–38, 133–147.
8. McCubbin MM, Milavetz G, Grandgeorge S, Weinberger M, Ahrens R, Sargent C, Vaughan LM. A bioassay for topical and systemic effect of three inhaled corticosteroids. *Clin Pharmacol Thera* 1995;57:455–460.
9. Stewart BS, Ahrens RC, Carrier S, Frosolono M, Lux C, Han S-H, Milavetz G. Demonstration of *in vivo* bioequivalence of a generic albuterol metered-dose inhaler to Ventolin. *Chest* 2000;117:714–721.
10. Ahrens RC, Hendeles L, Clarke WR, Dockhorn RJ, Hill MR, Vaughan LM, Lux C, Han S-H. Therapeutic equivalence of Spiros dry powder inhaler and Ventolin metered-dose inhaler. A bioassay using methacholine. *Am J Respir Crit Care Med* 1999;160:1238–1243.
11. Parameswaran KN, Inman MD, Ekholm BP, Morris MM, Summers E, O'Byrne PM, Hargreave FE. Protection against methacholine bronchoconstriction to assess relative potency of inhaled $beta_2$-agonist. *Am J Respir Crit Care Med* 1999;160:354–357.
12. Clark DJ, Lipworth BJ. Dose-response of inhaled drugs in asthma. An update. *Clin Pharmacokinet* 1997;32:58–74.
13. Barnes PJ, Pedersen S, Busse WW. Efficacy and safety of inhaled corticosteroids. New developments. *Am J Respir Crit Care Med* 1998;157 (3 Pt 2):S1–53.
14. Lipworth BJ. Clinical pharmacology of corticosteroids in bronchial asthma. *Pharmacol Ther* 1993;58:173–209.

15. Barnes PJ. Inhaled glucocorticoids for asthma. *N Engl J Med* 1995; 332:868–875.
16. Szefler SJ, Leung DYM. Glucocorticoid-resistant asthma: pathogenesis and clinical implications for management. *Eur Respir J* 1997;10:1640–1647.
17. American Thoracic Society. Committee on diagnostic standards for nontuberculosis respiratory disease: definition and classification of chronic bronchitis, asthma, and pulmonary emphysema. *Am Rev Respir Dis* 1962;85:762–768.
18. American Thoracic Society. Standardization of spirometry. 1994 update. *Am J Respir Crit Care Med* 1995;152:1107–1136.
19. Cockcroft DW, Killian DN, Mellon JJ, Hargreave FE. Bronchial reactivity to inhaled histamine: a method and clinical survey. *Clin Allergy* 1997;7:235–243.
20. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. New York: Chapman and Hall; 1993.
21. Little SA, Chalmers GW, MacLeod KJ, McSharry C, Thomson NC. Non-invasive markers of airway inflammation as predictors of oral steroid responsiveness in asthma. *Thorax* 2000;55:232–234.
22. van Rensen ELJ, Straathof KCM, Veselic-Charvat MA, Zwinderman AH, Bel EH, Sterk PJ. Effect of inhaled steroids on airway hyperresponsiveness, sputum eosinophils, and exhaled nitric oxide levels in patients with asthma. *Thorax* 1999;54:403–408.
23. Kips JC, O'Connor BJ, Inman MD, Svensson K, Pauwels RA, O'Byrne PM. A long-term study of the antiinflammatory effect of low-dose budesonide plus formoterol versus high-dose budesonide in asthma. *Am J Respir Crit Care Med* 2000;161:996–1001.
24. Taylor DA, Jensen MW, Kanabar V, Engelstatter R, Steinijans VW, Barnes PJ, O'Connor BJ. A dose-dependent effect of the novel inhaled corticosteroid ciclesonide on airway responsiveness to adenosine-5′-monophosphate in asthmatic patients. *Am J Respir Crit Care Med* 1999;160:237-243.
25. Swystun VA, Bhagat R, Kalra S, Jennings B, Cockcroft DW. Comparison of 3 different doses of budesonide and placebo on the early asthmatic response to inhaled allergen. *J Allergy Clin Immunol* 1998; 102:363–367.
26. Sont JK, Willems LNA, Bel EH, van Krieken JH, Vandenbroucke JP, Sterk PJ, and the AMPUL Study Group. The clinical control and histopathological outcome of asthma using airway hyperresponsiveness as an additional guide of long-term treatment. *Am J Respir Crit Care Med* 1999;159:1043–1051.
27. Jatakanon A, Kharitonov S, Lim S, Barnes PJ. Effect of differing doses of inhaled budesonide on markers of airway inflammation in patients with mild asthma. *Thorax* 1999;54:108–114.
28. Kraan J, Koeter GH, van der Mark TW, Boorsma M, Kukler J, Sluiter HJ, de Vries K. Dosage and time effects of inhaled budesonide on bronchial hyperreactivity. *Am Rev Respir Dis* 1988;137:44–48.
29. van Grunsven PM, van Schayck CP, Molema J, Akkermans RP, van Weel C. Effect of inhaled corticosteroids on bronchial responsiveness in patients with "corticosteroid naive" mild asthma: a meta-analysis. *Thorax* 1999;54:316–322.
30. Pedersen S, Hansen OR. Budesonide treatment of moderate and severe asthma in children: a dose-response study. *J Allergy Clin Immunol* 1995;95:29–33.
31. Fleiss JL. The Design and Analysis of Clinical Experiments. New York: John Wiley & Sons; 1999. p. 187–192.
32. Johansson SA, Dahl R. A double-blind dose-response study of budesonide by inhalation in patients with bronchial asthma. *Allergy* 1988; 43:173–178.
33. Chow S-C, Liu J-P. Design and Analysis of Bioavailability and Bioequivalence Studies. 2nd ed. New York: Marcel Dekker; 2000. p. 257–331.